

ABSTRACT OF THE DISCLOSURE

The present invention provides pitch conversion processing technology capable of minimizing the distortion of speech sound naturalness. A speech waveform in a pitch-unit is considered to be divided into two segments: 1) the segment of β , that starts from the minus peak, where the waveform depending on the shape of vocal tracts appears, and 2) the segment of γ where the waveform depending on the vocal tract shape is attenuating and converging on the next minus peak. In addition, α is the point where a minus peak appears along with the glottal closure. Based on characteristics of speech waveforms, the present invention processes waveform for converting pitch in the segment of γ just before the next minus peak, which is least affected by the minus peak associated with the glottal closure. As such, waveform processing can be performed by keeping the complete contour of waveform at around the peak, and thereby reducing the effects of pitch conversion.